## REMARKS

By this Amendment, Claims 3, 6, and 7 are amended and Claim 5 is canceled. The preamble of Claim 3 has been amended to include "constructing a model for predicting molecular behavior using marker molecules." Support for this amendment can be found, for example, on page 5, lines 10-11 of the specification.

Claim 3 has also been amended to include the step of "receiving data related to one or more chemical or biological properties of a set of reference molecules." Support for "receiving data" can be found on page 5, lines 10-31 and page 14, lines 3-15 of the specification. For example, on page 5, lines 10-31, an illustration is presented where protein binding is chosen as a property of interest. It is indicated that "a reference ('training') set of molecules is classified according to the property of interest." An exemplary set of ten molecules is presented with percent protein bound data. It is then indicated that the training set of molecules "can be classified into molecules that possess high protein binding and molecules that possess low protein binding." A table demonstrates the classification based on the percent protein bound data. In order for the classifying step to occur, the percent protein bound data must be provided and "received." Similarly, on page 14, lines 3-15, a "training dataset" composed of compounds listed in *Goodman & Gilman's The Pharmacological Basis of Therapeutics,* 9th ed. was used. This reference provided percent protein bound data. In order for the dataset to be composed and used in the described method, it had to be "received" from the *Goodman & Gilman* reference. Thus, support for a "receiving data" step is provided in the specification and the step is not new matter.

Support for "one or more chemical or biological properties" can be found, for example, on page 2, lines 21-23 of the specification and in original Claim 3. Specifically, page 2, lines 21-23 describe classifying a set of reference molecules "as either possessing or not possessing *at least one* property." Similarly, original Claim 3 called for classifying molecules as either possessing or not possessing "at least one property." The phrase "at least one property" indicates that the classification may be based on "one or more" properties. As indicated above, prior to classifying a set of reference molecules based on chemical or biological data, the data must be "received." Thus, in the "receiving" step of Claim 3, data relating to "one or more" chemical or

biological properties and received. This limitation is fully supported by the specification and the original claims and does not constitute new matter.

Claim 3 has also been amended to clarify that the "classifying" step classifies "the respective molecules" in the set of reference molecules for which data was received in the "receiving" step. This amendment merely clarifies that the classifying step involves classifying each individual molecule in the set of reference molecules rather than classifying the set as a whole into one class. This amendment is supported in the specification, for example, on page 5, lines 19-31. There it is indicated that a training set of molecules is "classified into molecules that possess high protein binding and molecules that possess low protein binding." An example is provided where ten molecules are classified, some classified as possessing high protein binding and some classified as not possessing high protein binding. Thus, each individual molecule in the set is classified rather than the entire set being assigned to the same class. Therefore, the amendment finds support in the specification and does not introduce new matter.

Other amendments to Claim 3 incorporate limitations previously found in Claim 5. The "choosing" step previously in Claim 5 has been amended to call for "choosing said first molecule as a marker molecule if said first molecule has a fractions-correctly-predicted metric which exceeds said threshold value for a pre-selected minimum distance." Previously, the "choosing" step called for "choosing, as said set of marker molecules, those molecules of said subset having a fractions-correctly-predicted metric which exceeds said threshold value for a pre-selected minimum distance." Thus, the only difference between the two forms of the "choosing" step is "choosing said first molecule" as opposed to "choosing...those molecules." Such a difference is fully supported by both the specification and the original claims. In describing "one specific method" for selecting marker molecules, the specification provides on page 10, lines 3-5, that "a set of marker molecules is defined as every DTC molecule having a MOLCNT of equal to or greater than a selected value while maintaining a selected minimum FCP threshold." In order to define a set of marker molecules, individual "DTC molecule[s]" must be separately evaluated to determine whether they are to be part of a set of marker molecules. Thus, in this example, the "choosing" of marker molecules requires an evaluation of MOLCNT and FCP values one molecule at a time. The "choosing" step of Claim 3 is directed to at least one such evaluation. Thus, "choosing said first molecule" is supported by the specification and is not new matter. Further support can be found on page 10, lines 11-14 and page 11, lines 13-16 of the

specification, which provide an example of a "marker molecule set containing only Naproxen" and block 18 of Figure 1, which indicates that "*one* or more marker molecules" may be selected "from the subset." These selections indicate that a marker molecule set may contain only a single molecule and further support the Claim 3 limitation calling for "choosing said first molecule."

Claim 6 has been amended so that it no longer depends on canceled Claim 5. Claim 7 has been amended so that it depends on Claim 3 and to incorporate limitations previously found in Claim 5. The recitation of repeating "for other molecules of said subset" in Claim 7 finds support, for example, in block 18 of Figure 1 where "one or more marker molecules from the subset" may be selected. If more than one marker molecule is to be selected, repetition of the steps in Claim 3 is required. However, it is not required in the general embodiment described on page 5, line 10 to page 6, line 15 and Figure 1 that all molecules of the subset be considered. Rather, "one or more" of the subset molecules may be selected. As such, there is support in the specification for the Claim 7 limitation "other molecules of said subset." Claims 3-4 and 6-8 remain pending in this application.

Objection

The Examiner has objected to Claim 3 because an "e" was missing from the word "one" as originally filed. The Applicants respectfully submit that the present amendment adding an "e" overcomes the objection.

Rejections Under § 101

The examiner has rejected Claims 3-8 as being directed to non-statutory subject matter. The Applicants respectfully submit that, as amended, Claims 3-4 and 6-8 are directed to statutory subject matter because they disclose manipulating data representing physical objects.

Claim 3 has been amended to recite "receiving data related to one or more chemical or biological properties of a set of reference molecules." The chemical and biological properties of the reference molecules are measurements of physical objects. Such data must be obtained from the physical world. These are real molecules that exist and have been characterized. In other words, if Claim 3 is being implemented as an algorithm in a computer system, the algorithm itself cannot calculate the "one or more chemical or biological properties." For example, at the bottom of page 5 in the specification, a sample set of reference molecules is provided along with

percent protein bound data. The property of " % protein bound" comes from the physical world and is not calculated by a computer algorithm.

Implicit in Claim 3 is that real world experimentation determined the claimed chemical and biological properties. The specification provides an example on page 14. There, data related to the chemical property of interest (percent protein bound) was obtained from *Goodman & Gilman's The Pharmacological Basis of Therapeutics,* 9th ed. Thus, the data used in the example originated from real world experimentation. As such, Claim 3 is analogous to the method claim allowed in *In re Abele*, 684 F.2d 902 (CCPA 1982). In that case, the method claim was directed to an algorithm that used "X-ray attenuation data produced in a two dimensional field by a computed tomography scanner." *Id.* at 908. The CCPA held that the claim was directed to statutory subject matter because the data used by the algorithm was obtained from a physical process. The claim itself did not require the step of physically measuring the X-ray attenuation data. Rather, the claim required that the data had been "produced...by a computed tomography scanner." Therefore, a claimed algorithm may use data input by a user, input directly from a database, or input by any other means so long as the data was originally obtained from a physical measurement. The actual step of making the physical measurement need not be a claim limitation. Thus, in the step in Claim 3 requiring "receiving data related to one or more chemical or biological properties of a set of reference molecules," the data may be received from any source (e.g., a reference article or book, a computer database, an interface with experimental instrumentation) as long as the data was originally obtained by a physical measurement. As illustrated by *In re Abele*, such a claimed process is directed to statutory subject matter.

The applicants respectfully submit that Claims 3-4 and 6-8 require the "manipulation of data representing physical objects" and therefore are directed to statutory subject matter (M.P.E.P. § 2106(IV)(B)(2)(b)(i)).

Rejections Under § 112

The Examiner has rejected Claims 5-8 as lacking support in the specification for the limitation of "repeating...for at least some other molecules of said subset." The Applicants have canceled Claim 5. It is respectfully submitted that Claims 6-8 are allowable.

Rejections Under § 102

The Examiner has rejected Claims 3-8 as allegedly being anticipated by Stanton et al. The Applicants respectfully submit that Stanton et al. does not anticipate any of Claims 3-4 and 6-8 as amended. With respect to independent Claim 3, at a minimum, Stanton et al. does not disclose the steps of "defining...a fractions-correctly-predicted metric," "counting the number of molecules...," and "choosing said first molecule as a marker molecule..."

Claim 3 further requires "sorting" molecules "in descending order of numerical similarity" from a "first molecule." The Examiner has alleged that the squared Euclidean distance similarity metric disclosed in Stanton et al. reads on the "similarity metric" of now amended Claim 3 (previously a limitation in Claim 5; see 9/22/03 Office Action, page 7). In addition, the Examiner has alleged that the disclosure in Stanton et al. of sorting molecules in descending order of numerical similarity based on Euclidean distance reads on the "sorting" step in now amended Claim 3 (previously a limitation in Claim 5; see 9/22/03 Office Action, page 8). The Examiner goes on to allege that the last column of Table 2 in Stanton et al., titled "hit rate as a % of compds tested," reads on the "fractions-correctly-predicted metric" in now amended Claim 3 (previously a limitation in Claim 5; see 9/22/03 Office Action, pages 8-9).

The Applicants respectfully disagrees with the Examiner's analysis. Claim 3 explicitly defines the "fractions-correctly-predicted metric as the number of molecules in said range which are also members of said subset divided by the total number of molecules in said range." The "range" in this limitation is the "range in molecules of similarity distance away from" the "first molecule." In contrast, Stanton et al. does not disclose similarity distances in Table 2, let alone a range of similarity distances. Instead, Table 2 discloses the hit percent for *all* compounds returned in the nearest neighbor search for each of the eleven disclosed examples. Figure 5b plots the hit percent for the *combination* of molecules from *all* eleven examples as a function of nearest neighbor Euclidean distance; however, Figure 5b does not disclose the hit percent as a function of nearest neighbor distance for each independent example. In other words, Stanton et al. does not disclose, nor is it possible to determine from Table 2 or Figure 5b, hit percents of the nearest neighbors to *one* of the eleven particular original queries ("a first molecule"). Furthermore, the hit percent in Stanton et al. is the hit percent for the molecules tested that have a *given nearest neighbor distance* from the original query. Figure 5b does not report, nor is it disclosed elsewhere in Stanton et al., the hit percent for *all* molecules within the *range* between

the original query and a given nearest neighbor distance. Thus, Stanton et al. does not disclose the "fraction-correctly-predicted metric" as defined in Claim 3, and therefore does not anticipate Claim 3.

Moreover, Stanton et al. does not disclose the step of "counting the number of molecules away from said first molecule at which the fractions-correctly-predicted metric for said first molecule drops below a threshold value." The Examiner has pointed out that Stanton et al. discloses that the 20-30 closest neighbors to a given query were selected for subsequent purchase and screening (see 9/22/03 Office Action, page 8). However, merely selecting a number of closest neighbors without a specified criteria for the selection cannot meet the limitations of Claim 3, which require counting the number of molecules when "the fractions-correctly-predicted metric for said first molecule drops below a threshold value." In contrast to this specified criteria, Stanton et al. states with respect to the 20-30 closest neighbors that "[i]nitially, this number was chosen arbitrarily" (page 22, col. 1, line 6). Furthermore, Stanton et al. merely discloses selecting nearest neighbors and does not disclose "counting." The Examiner has also pointed out that Stanton et al. discloses that a 20 percent hit rate may be obtained by considering only the available compounds within a nearest neighbor distance of 1.9 or less (9/22/03 Office Action, page 9; Stanton et al., page 26, col. 2, lines 9-11). However, as discussed above, the hit rate in Stanton et al. is not the same as the claimed "fractions-correctly-predicted metric." Furthermore, Stanton et al. does not call for counting the number of molecules within a nearest distance of 1.9 to the original query, nor is that number provided. Indeed, it would not be possible to determine the count from the information plotted in Figure 5, because Figure 5 combines data from *all* 11 examples in Table 2. As such, counts of molecules at given nearest neighbor distances from a particular original query ("a first molecule") are not plotted. Thus, Stanton et al. does not disclose the "counting" step of Claim 3 and, for this additional reason, does not anticipate Claim 3.

Stanton et al. also does not disclose the limitation of the "choosing" step in Claim 3. Claim 3 requires "choosing said first molecule as a marker molecule if said first molecule has a fractions-correctly-predicted metric which exceeds said threshold value for a pre-selected minimum distance." Stanton et al. discloses calculating hit percents for nearest neighbor searches on original queries (Stanton et al, Table 2 and Figure 5b). As discussed above, these hit percents are not the same as the "fractions-correctly-predicted metric." Furthermore, Stanton et

al. does not disclose choosing particular original queries based on the hit percents for a pre-selected minimum distance. The disclosure in Stanton et al., of a set of original query examples for which the *combined* data indicated a 20% hit rate within a nearest neighbor distance of 1.9, was not used as a criteria for choosing certain of the original queries over others. Since the "choosing" step is not at all disclosed in Stanton et al., for this additional reason Claim 3 is not anticipated by Stanton et al.

It is also worth noting the difference between Claim 3 as a whole and Stanton et al. The Examiner has alleged that finding hits within a large combinatorial library as disclosed in Stanton et al. reads on the "classifying" step of Claim 3 (9/22/03 Office Action, pages 6-7). Thus, the Examiner has argued that the large combinatorial library disclosed in Stanton et al. reads on the "set of reference molecules" of Claim 3. Furthermore, the Examiner has alleged that performing hierarchical cluster analysis to obtain subsets of hits as disclosed in Stanton et al. reads on the "subset" of Claim 3. If these allegations were correct, Claim 3 would require, among other steps, "selecting a first molecule" from a "subset" created in the hierarchical cluster analysis and sorting all of the molecules from the "set" of the large combinatorial library "in descending order of numerical similarity." However, in Stanton et al., Figure 5 only depicts Euclidean distances from original query compounds for the compounds returned in nearest neighbor searches, not the entire combinatorial library of compounds. Furthermore, as discussed above, Stanton et al. does not disclose choosing particular original queries used in nearest neighbor searches. Thus, Claim 3 as a whole is not disclosed in Stanton et al. See M.P.E.P. § 2106(II)(C): "[W]hen evaluating the scope of a claim, every limitation in the claim must be considered. Office personnel may not dissect a claimed invention into discrete elements and then evaluate the elements in isolation. Instead, the claim as a whole must be considered." (emphasis in original).

Claim 4, 6, and 7 depend from Claim 3, and Claim 8 depends from Claim 7. Therefore, for at least the same reasons as Claim 3, Claims 4 and 6-8 are not anticipated by Stanton et al. Furthermore, Claim 7 requires a "repeating" step using "a plurality of different threshold values and minimum distances." Stanton et al. does not disclose the use of multiple threshold values for a "fractions-correctly-predicted metric"; nor does it disclose the use of multiple minimum nearest neighbor distances. Thus, for this additional reason, Stanton et al. does not anticipate Claims 7 and 8.

## CONCLUSION

The Applicants have amended the claims in order to address the Examiner's objection and rejections based on sections 101, 112, and 102. The applicants respectfully submit that all of the rejections have been overcome and the cited art distinguished. As such, allowance of all pending claims is respectfully requested.

If the Examiner has any questions that may be answered by telephone, she is invited to call the undersigned directly.

Please charge any additional fees, including any fees for additional extension of time, or credit overpayment to Deposit Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: 2/10/04

By: _____

Thomas R. Arno
Registration No. 40,490
Attorney of Record
Customer No. 20,995
(619) 235-8550

S:\DOCS\REM\REM-1138.DOC
021004